View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Springer - Publisher Connector

Appl Geomat (2009) 1:97–110
DOI 10.1007/s12518-009-0013-1

ORIGINAL PAPER

A new expert system module for building detection in urban areas using spectral information and LIDAR data

Ayman Rashad Elshehaby · Lamyaa Gamal El-deen Taha

Received: 16 May 2009 / Accepted: 13 November 2009 / Published online: 9 December 2009
© Società Italiana di Fotogrammetria e Topografia (SIFET) 2009

Abstract Buildings are one of the most important geospatial features for spatial analysis and mapping. Building extraction has been an active research topic in computer vision as well as digital photogrammetry in recent years. Building detection is the process of obtaining the approximate position and shape of a building, while building extraction can be defined as the problem of precisely determining the building outlines, which is one of the critical problems in digital photogrammetry. Building information is extremely important for many applications such as urban planning, telecommunication, three-dimensional city modeling, or extraction of unauthorized buildings over agricultural lands. Three approaches for building detection based on maximum likelihood classification have been compared, firstly, building detection from classification of multispectral satellite image only. The second approach is building detection from classification of multispectral satellite image, while the height information from Light Detection and Ranging (LIDAR) data is applied as an additional channel together with spectral channel. The third approach is building detection based on classification of multispectral satellite image where normalized difference vegetation index (NDVI) and the height information from LIDAR data are applied as additional channels together with spectral channel. The contributions of the individual cues used in the classification have been evaluated. The three approaches were tested using urban blocks containing different sizes, roof color and shapes of buildings. The results show that the third approach is the best for building detection followed by the second approach then the first approach. The third approach appears to be quite successful especially in solving the problem of building detection for those urban blocks that contain closely located buildings as well as in separation of buildings from trees. The third approach results have been improved by developing a building detection module based on integration of classified image, elevation data, and spectral information. A rule-based expert system consists of essentially hypothesis (output; buildings), and variables of a knowledge base were developed in the knowledge engineer of ERDAS Imagine for post-classification refinement of initially classified output building mask. Classification rules were enriched with ancillary data such as the normalized digital surface model and the NDVI. Each rule is a representation of each node in the tree that describes a building class or probability of presence of buildings pixel. Then, the building detection result has been evaluated. It has been found that the use of an expert system, which considers expert knowledge, would further help in the discrimination of the classes and improve classification accuracy of buildings. The overall accuracy of expert classification was 96% and kappa coefficient was 0.95.

Keywords Multi-cue integration · Building detection · Building extraction · nDSM · NDVI · LIDAR · Digital photogrammetry · Computer vision · Knowledge-based expert system · Hierarchical decision tree classifier

A. R. Elshehaby
Banha University,
Cairo, Egypt

L. G. E.-d. Taha (✉)
National Authority of Remote Sensing and Space Science (NARSS),
Cairo, Egypt
e-mail: Lamyaa@narss.sci.eg

Introduction

The recent availability of commercial high-resolution satellite imaging sensors such as IKONOS provides a new data source for building extraction. High spatial resolution of the imagery specifies very fine details in urban areas and facilitates the classification and extraction of urban-related features such as roads and buildings.

Since manual extraction of features from imagery is a very slow process; automated methods have been proposed to improve the speed. During the past years, numerous classification algorithms have been developed. They can be divided into unsupervised and supervised approaches. Most of the recent work on building extraction from high-resolution satellite images is based on supervised techniques. In a supervised classification, two basic steps are carried out. First, in a training stage, an operator digitizes training areas that describe typical spectral and textural characteristics of the dataset. In the following classification stage, each pixel of the dataset is assigned to a land cover class. For this classification stage, a lot of different approaches such as minimum distance, Mahalanobis, or maximum likelihood classification are available.

Spectral information has been widely used as a data source for thematic mapping applications (Haala and Walter 1999). A common goal during data acquisition in built-up areas is the detection of objects like streets and buildings. However, this can be difficult if only spectral information is used, since, for some areas, roofs and streets are built of very similar material. This complicates or even prevents the discrimination of these objects due to their similar reflectance (Haala and Walter 1999). For this reason, a multi-cue integration of remotely sensed data can be used for solving of this problem.

An increased number of cues are derived from remotely sensed data. An important point not only for higher success rate but also lower processing costs is the number and type of used cues for object extraction. Choosing correct cue combination can help us for feature extraction (Baltsavias 2002). Combining of elevation data and spectral information for building extraction is quite promising. Since height data has been approved as a very valuable information for raised objects discrimination (Guo and Yasuoka 2002b).

In digital photogrammetry for elevation determination, stereo image matching techniques determine corresponding pixels or features in two overlapping images. Conventional image matching techniques only supply a digital surface model (DSM). This means that matching occurs on the top of man made objects such as buildings, or on the top of the vegetation rather than the terrain surface and hence does not represent the terrain surface (Lu et al. 2003). One can use this DSM associated with a digital terrain model (DTM) resulted from topographic maps.

In light detection and ranging (LIDAR) technology, a LIDAR sensor system permits an aircraft flyover to quickly collect a height for large regions with high vertical accuracy and high point density. LIDAR can collect three-dimensional points from both first and last returns. The LIDAR points being on the terrain are separated from points on buildings and other object classes; DSM and DTM can be computed (Rottensteiner and Briese 2002).

With the assistance of the expert system, it is possible to integrate multi-cue derived from remotely sensed data. Knowledge-based expert systems continue to be used extensively in remote sensing research. Currently, researchers are using knowledge-based rule image analysis techniques to encode rules used by human interpreters, which can be used by a computer for feature extraction (Forghani 1999).

Muchoney et al. (2000) used a decision tree classifier to extract land cover information from MODIS data. Tso and Mather (2001) summarized numerous applications of hierarchical decision tree classifiers–the most general type of knowledge-based classifier. Pal and Mather (2003) assessed the effectiveness of decision tree methods for land cover classification.

In this paper, three approaches for building detection based on maximum likelihood classification have been compared. Building detection has been performed by classification of multispectral image only. Also by combining laser data and color imagery in a single classification step resulting in a multichannel classification (multispectral+ normalized DSM (nDSM)) and (multispectral+normalized difference vegetation index (NDVI)+nDSM). Although some promising results from the third approach of building detection has been achieved, so far, it still needs some improvement. An expert system for post-classification refinement has been implemented using knowledge engineer that's available in ERDAS Imagine 8.7 Software.

The building extraction process is composed of two basic steps: (1) building detection and (2) building (delineation) extraction.

Study area

The study area, comprising 1 km^2, was chosen at Al-Sayda Zeinab, Cairo, Egypt, which is covered by sheet ط15, scale 1:5,000 produced from aerial photography in 1978, and was revised in 2006.This area is located in middle of Cairo city and has a very representative urban scene.

Data sources

1. Multispectral IKONOS image (1 m resolution (fused)) which has been processed using ERDAS Imagine 8.7 (see Fig. 1): IKONOS image has been used instead of RGBI because RGBI is not available in our hands.
2. Laser scanner data captured with TopoSys scanner contains the first and the last echoes of the laser beam.

According to the specification of laser scanner, it delivers very high point densities, 83,000 measurements per second; the average measurement density is 3 measurements/m^2; the vertical accuracy of LIDAR data is 15 cm; and the horizontal accuracy is 50 cm. Up to 1999, the Toposys instrument isn't able to measure the reflected signal intensity, so it gives pure geometric data, and it can acquire first pulse or last pulse alternatively. Figure 2 shows digital elevation model derived from LIDAR.

3. Check points have been measured using differential global positioning system (GPS) with an accuracy of 5 cm in x, y, and 3 cm in z. Twenty points in the bare earth and twenty points above buildings.
4. Large-scale planimetric map 1:5,000 (2006) was produced from aerial photos (Date 1977–1978) and updated in 2006. The map is published by Egyptian Surveying Authority (ESA) in Egyptian Transverse Mercator projection.
5. Multispectral QuickBird image 2007 of the same area obtained from Google Earth and processed using ERDAS Imagine 8.7 (see Fig. 3): this image was used for revision of building shapes instead of field revision because there is a large shadow area in IKONOS image.

Fig. 3 QuickBird image obtained from Google Earth.

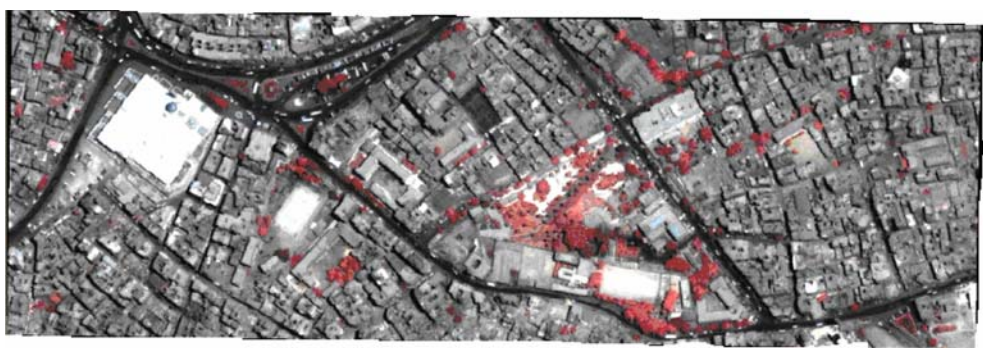

Fig. 1 Multispectral IKONOS image.

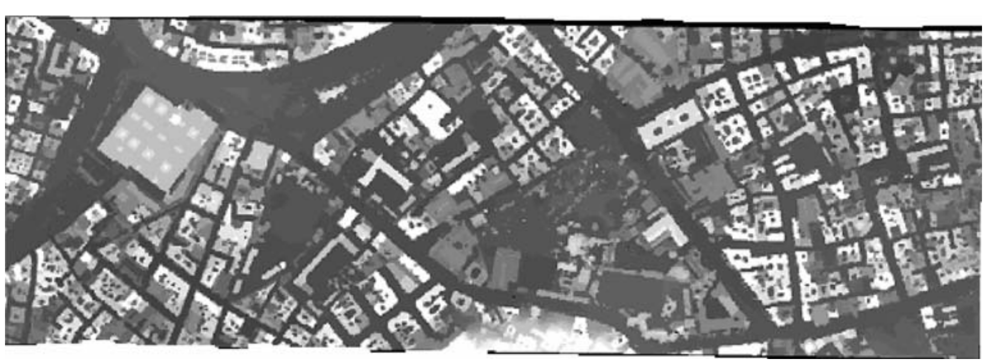

Fig. 2 Light detection and ranging data (LIDAR data).

Methodology

- Map scanning, georeferencing, and vectorization.
- Creation of LIDAR DSM and DTM.
- Calculation of nDSM.
- IKONOS image has been orthorectified using LIDAR DSM.
- Calculation of NDVI.
- Maximum likelihood classifier has been used for building detection, firstly, building detection from classification of multispectral satellite image only. The second approach is building detection from classification of multispectral satellite image, while the height information from LIDAR data is applied as an additional channel together with spectral channel. The third approach is building detection based on classification of multispectral satellite image where NDVI and the height information from LIDAR data are applied as additional channels together with spectral channel. Signatures has been collected and evaluated from the resulted three cases. Accuracy assessment of classifications was carried out using overall accuracy and kappa coefficient. Seventy randomly selected points were used for this purpose.
- Then, in each of the resulting three classifications, buildings were separated in a mask.
- Morphological opening with kernel size of 3×3 followed by morphological closing with kernel size of 3×3 have been applied to each of the resulted three building masks using ENVI 4.2 software. Accuracy assessment was carried out using overall accuracy and kappa coefficient. Seventy randomly selected points were used for this purpose.
- The third approach result (buildings mask) has been improved by developing a building detection module based on integration of classified image, elevation data (LIDAR data nDSM), and spectral information derivative (NDVI) using knowledge engineer for post-classification refinement of initially resulted building mask.
- Rectified multispectral QuickBird image has been used for revision of the resulted buildings especially buildings that are not discernible enough due to shadows of IKONOS image (see Fig. 3).
- This step was used instead of the fieldwork.

Map scanning, georeferencing, and vectorization

Scanning is a very common procedure for transforming hardcopy maps into a digital format, where the output is a raster map.

Large-scale planimetric map of 1:5,000 was scanned with a scanner of 400 dpi then the map has been georeferenced using four points. After that, Scan 2CAD program has been used for automatic map vectorization.

Creation of LIDAR digital surface model (DSM) and digital terrain model (DTM)

The LIDAR points being on the terrain are separated from points on buildings and other object classes, a DTM and DSM can be computed (Rottensteiner and Briese 2002).

Preprocessing

The flight path has been calculated combining the GPS data from the aircraft and the data from the reference station (DGPS solution) by using Applanix POSGPS software. Then, the data from the measurements of the inertial measurement unit (IMU) must be integrated. Therefore, the Applanix POSPROC has been used. Afterwards, the position and orientation of the sensor system have been available. This data and the laser scanning data have been combined with the TopPIT software.

Post-processing

TopPIT Software package from TopoSys is used for processing of laser scanner data. This step contains the calculation of ground points from the combined position and laser file. The first echo has been used to generate a DSM; the calculated ground points are sorted into a regular grid, where one height value belongs to one pixel. With the TopPIT software, it is possible to modify the grid spacing of the output DSM (raster data DSM). 0.5 m grid spacing is chosen as elevation grid.

After that, DTM has been derived from the DSM using “bevefil” module in the TopoSys software, which erodes objects. Therefore, the objects contained in a DSM like vegetation and buildings are eliminated. Bevefil module is a bisectional filter (bisectional algorithm), which constructs a convex/concave covering, which is selected from the bottom to the top. Also, a median filter has been applied.

GPS check points data were measured with an accuracy of 5 cm in x, y and 3 cm in z. Twenty points in the bare earth and twenty points above buildings were used to check the accuracy of the DTM and DSM. While the accuracy of the generated DTM was computed to be 0.2 m, the accuracy of the DSM was found to be 0.28 m.

Calculation of normalized digital surface model (nDSM)

After computation of both DTM and DSM, nDSM has been calculated by subtraction of DTM from DSM (DSM–DTM) (Rottensteiner and Briese 2002; Haala 1999). The basic idea of using height data in a building extraction is that man made objects with different heights over the terrain can be detected by applying a threshold to the nDSM. Those areas of nDSM that fall above the userdefined threshold are considered to represent the three-dimensional objects (San and Turker 2007). The above ground features were separated from the terrain by applying a height threshold of 3.5 m to the nDSM since the heights of the buildings in our study area have been assumed higher than 3.5 m, so that an initial building mask is created which still contains vegetation and other objects.

Preprocessing of IKONOS image

IKONOS image has been orthorectified in order to remove the relief displacement using LIDAR DSM. Ten well-distributed map control points have been identified in both large-scale map and IKONOS image. The coordinates of these points were matched. First order polynomial and nearest neighbor resampling were used. The root mean square (RMS) in the east, the RMS in the north, and the total RMS error was 1.19, 0.83, and 1.45 m, respectively. Twelve well-distributed map control points have been identified in both georeferenced 1:5,000 map and IKONOS image and used as checkpoints. The RMS in the east, the RMS in the north, and the total RMS error was 1.24, 0.63, and 1.39 m, respectively.

Calculation of normalized difference vegetation index (NDVI)

The NDVI can be used to transform the multispectral data into a single image band representing vegetation. The NDVI values indicate the amount of green vegetation present in the pixel (Lu et al. 2003).

The NDVI can be calculated as follows: $\mathrm{NDVI} = (\mathrm{IR} - R)/(\mathrm{IR} + R)$, where

IR Near-infrared reflectance value
R Visible red reflectance value (Guo and Yasuoka 2002a).

The NDVI has been calculated using the red and near-infrared bands of the orthorectified IKONOS image. Figure 4 indicates the NDVI; one can see that the vegetation is represented by white color, and other objects are represented by the shade of gray. Figure 5 indicates the histogram of NDVI.

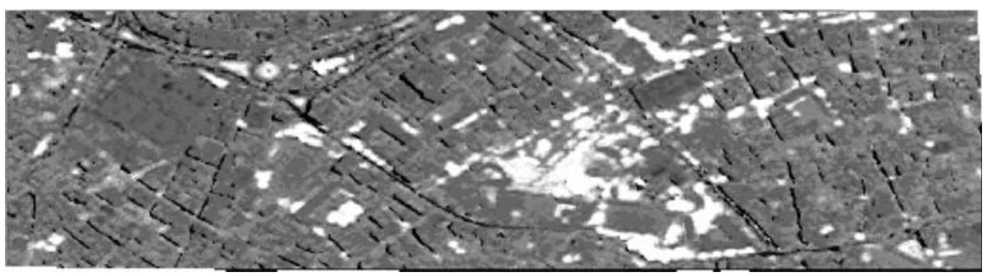

Fig. 4 Normalized difference vegetation index.

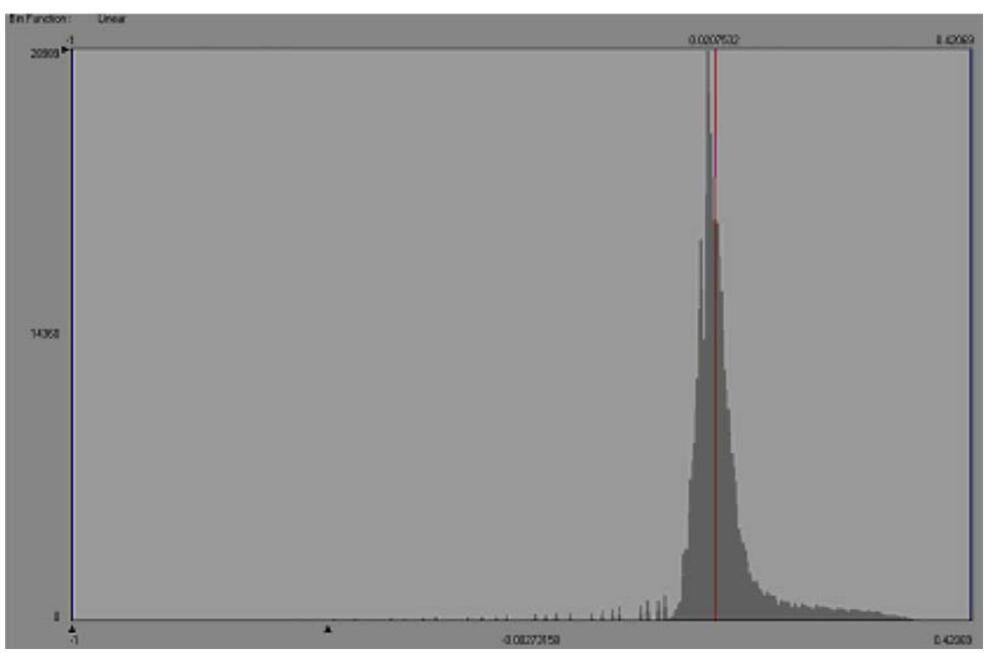

Fig. 5 Histogram of normalized difference vegetation index.

Building detection

Building detection based on maximum likelihood classification

Classification is the process of sorting all the pixels in an image into a finite number of individual classes.

Maximum likelihood classification is still the most widely used supervised classification algorithm. This method assumes that the probability distributions for the all input classes possess a multivariate normal distribution (Jensen 2005).

Three approaches for building detection based on maximum likelihood classification have been compared in separating the buildings from other classes.

In the first approach, the building has been detected from classification of multispectral satellite image only, height data only used for orthorectification before classification. The second approach is building detection from image classification where the information on the local height above the terrain (nDSM) is applied as an additional channel together with spectral channel (Fig. 6). The third approach is building detection based on integration of elevation data (nDSM) and spectral information (multispectral satellite image, NDVI), which means insertion of nDSM and NDVI as additional channels (Fig. 7).

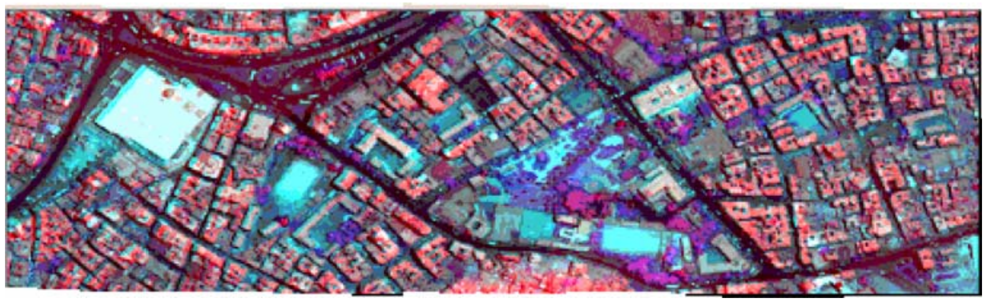

Fig. 6 False color composite of multispectral image with normalized digital surface model as an additional layer.

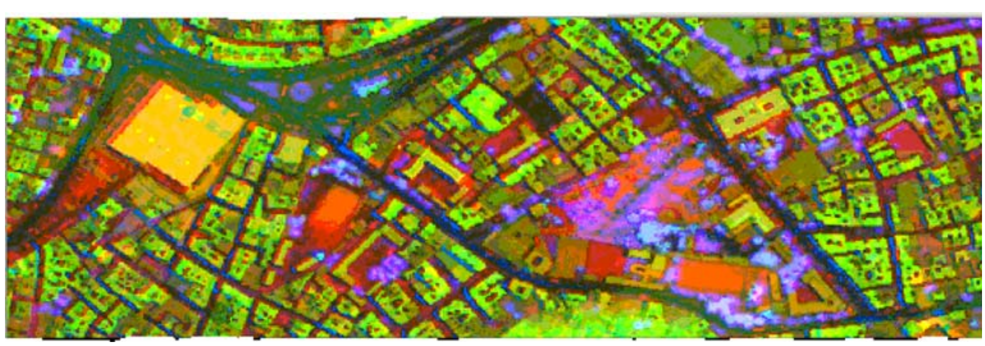

Fig. 7 Multispectral image with normalized digital surface model and normalized difference vegetation index as an additional layers.

Collection of spectral signatures Signatures collection is the first step in the classification process. Three classes were selected to represent the land use/land cover classes of the study area: buildings, roads, and vegetation. Thirty signatures have been collected in each class.

Signatures evaluation The objective of signatures evaluation is to ensure that they represent unique land covers and that they will produce the most accurate classification.

The collected signatures were evaluated, and the result is accepted before the classification process. An example on signatures evaluation has been given in Fig. 8, which indicates the histogram of the collected signatures for the first approach.

Building mask Then, in each of the resulting three classifications, buildings were separated in a mask.

Figure 9 shows an example of buildings mask resulted from the third approach; one can see some artifacts such as sparsely distributed pixels.

Analyses of causes of false detection are as follows:

1. The complexity of urban scene due to high density of buildings: the distribution of buildings ranges from sparse to very close.
2. Built-up areas suffer from problems due to occlusions and height discontinuities.
3. The quality of NDVI or nDSM.
4. Some areas contain a lot of trees as well, from individual trees to tree crowds; some trees are very close to buildings.
5. There are large areas covered by shadow in IKONOS images.
6. The classification accuracy depends on the building size. It decreases with the small building size. Some buildings have some parts missing (not completely detected).
7. Different image intensity for different buildings.

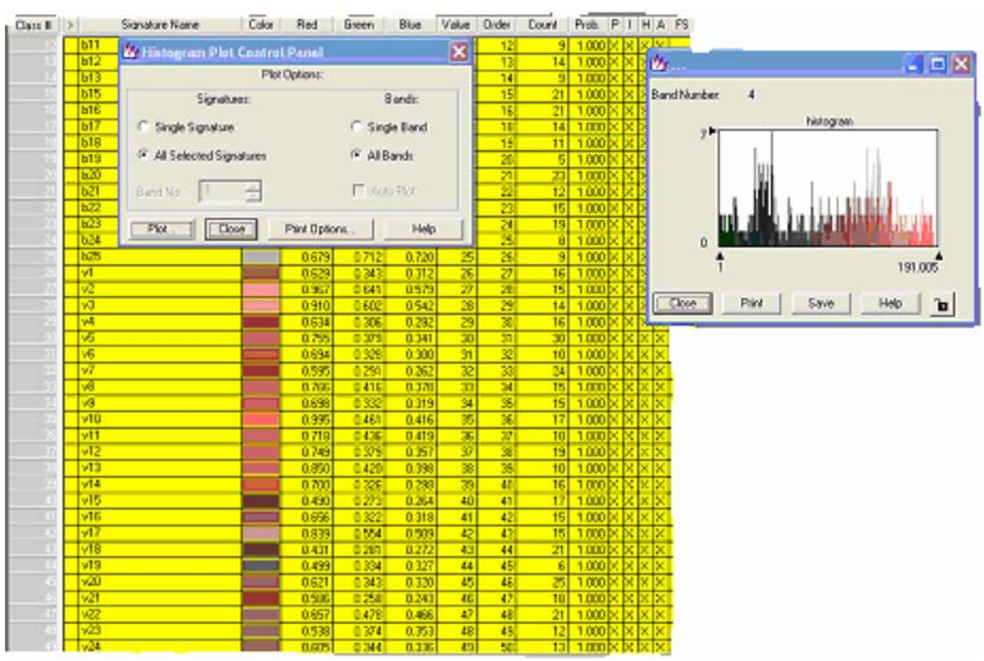

Fig. 8 Example of signature evaluation of multispectral classification.

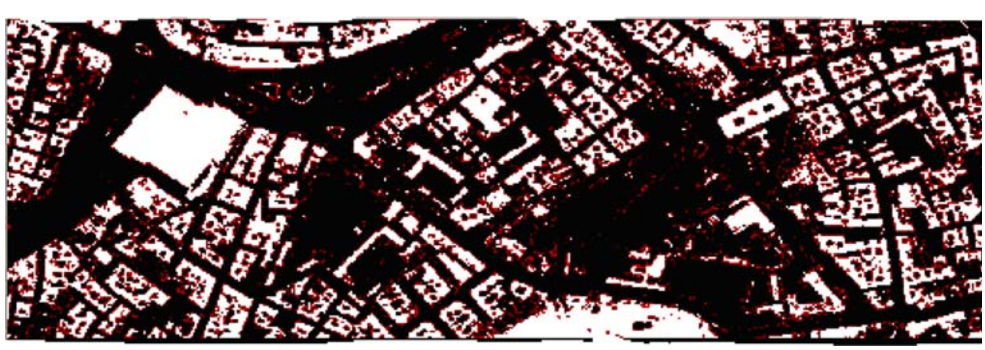

Fig. 9 Buildings mask resulted from the third approach.

Morphological operators Building mask may contain artifacts. In order to remove these artifacts, the opening and closing morphological operations were used (San and Turker 2007).

A morphological opening filter using a small (3×3) square structural element is to be applied to the initial building mask followed by morphological closing filter in order to erase small elongated objects such as fences and to separate regions just bridged by a thin line of pixels (Rottensteiner and Briese 2002). Figure 10 indicates an example of buildings resulted from the third approach after applying morphological operations. By comparing Fig. 9 (before applying morphological operations) and Fig. 10 (after applying morphological operations), one can see that the artifacts have been removed resulting in an improvement in the building detection results.

Figure 11 shows the workflow of the three approaches for building detection.

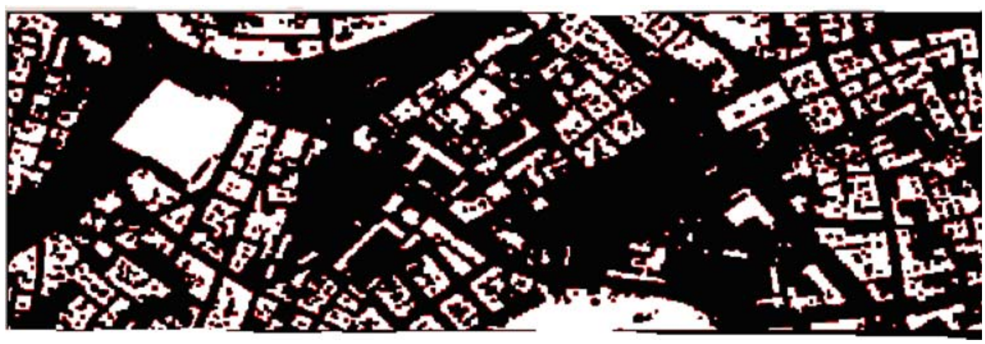

Fig. 10 Buildings mask resulted from the third approach after applying morphological operations.

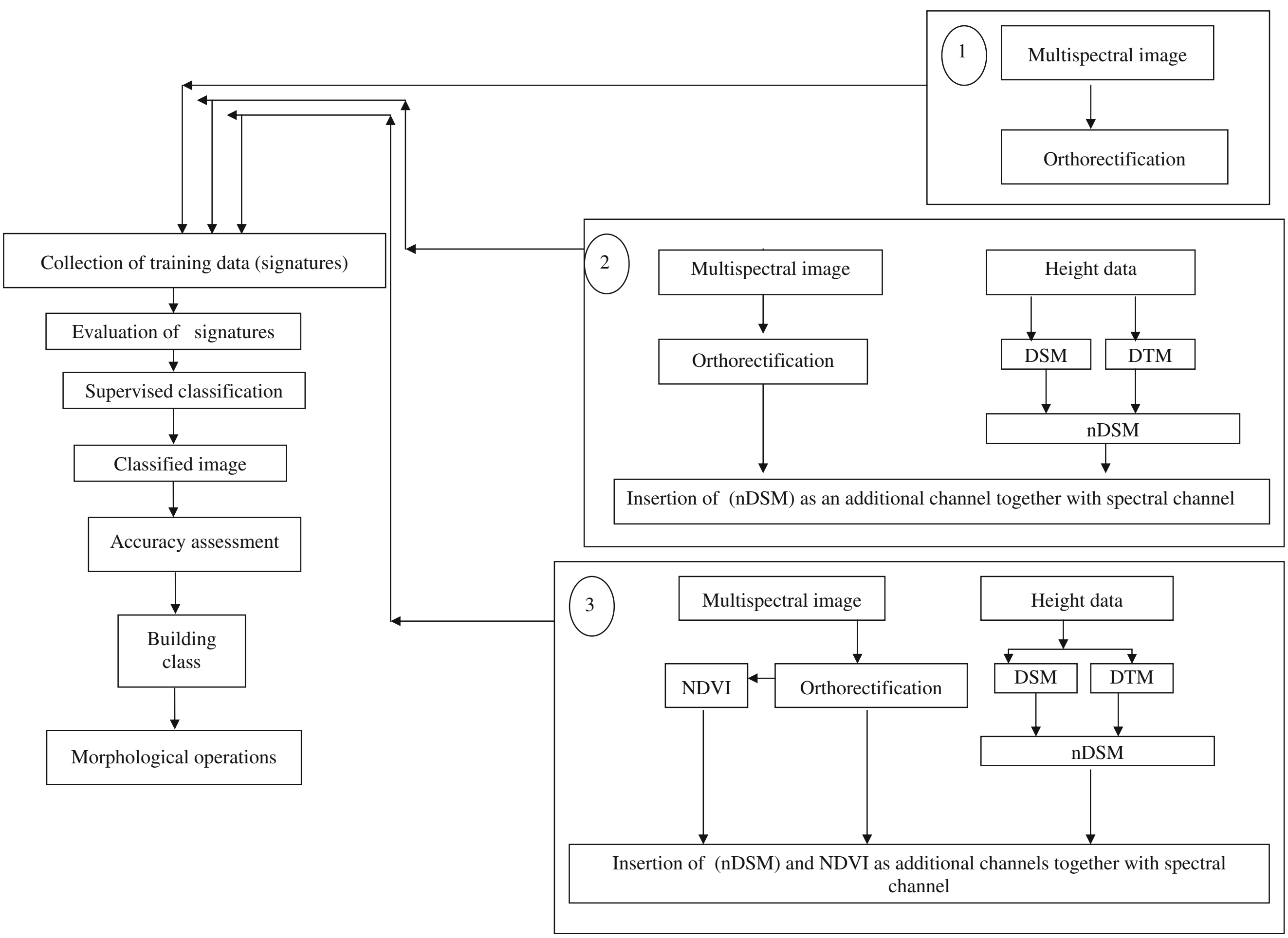

Fig. 11 The workflow of building detection from the three approaches.

Building detection based on expert system

The expert system makes use of layers of raster data, each layer relating to a type of "evidence" for the existence of a certain class (Nangendo et al. 2007).

Basic idea The NDVI and DSM are two key parameters, which define the difference between vegetated and non-vegetated objects (Lu et al. 2003). The buildings were differentiated from the trees using the previously calculated NDVI image by simply masking out the vegetated areas (San and Turker 2007). Here, classification is conducted based on such a simple fact: the objects, which have the height above a certain value, must be either trees or buildings; meanwhile, trees have high NDVI value, and NDVI of buildings is low. Similarly, grasslands or cultivated areas have low height (similar to terrain surface) but high NDVI; bare lands have low height, medium NDVI, and streets have low height and low NDVI (Guo and Yasuoka 2002a; Lu et al. 2003).

If the image consisted of five land covers: building, street, bare land, grassland, and tree. The general rule for this segmentation is quite simple but efficient as shown in Table 1, which indicates that building class will be found when nDSM is high and NDVI is low, street class will be found when nDSM is low and NDVI is low, bare land class will be found when nDSM is low and NDVI is medium, grassland class will be found when nDSM is low and NDVI is high, and finally, tree will be found when nDSM is high and NDVI is very high.

Table 1 Segmentation rule for urban objects.

Objects	NDVI	nDSM
Building	Low	High
Street	Low	Low
Bare land	Medium	Low
Grassland	High	Low
Tree	Very high	High

Table 2 indicates parameters for classification based on NDVI and nDSM. From the histogram of the NDVI, one can get the ranges of NDVI values at which different class appear. Building class will be found when the NDVI is less than −0.02 and the nDSM is greater than 3.5 m, street class will be found when the NDVI is less than −0.02 and the nDSM is less than 3.5 m, bare land class will be found when the NDVI ranges from −0.02 to 0.05 and the nDSM is less than 3.5 m, grassland class will be found when the NDVI ranges from 0.05 to 0.1 and the nDSM is less than 3.5 m, and finally, tree will be found when the NDVI is greater than 0.1 and the nDSM is greater than 3.5 m.

Table 2 Parameters for classification based on normalized difference vegetation index and a normalized digital surface model.

Objects	NDVI	nDSM
Building	<−0.02	>3.5
Street	<−0.02	<3.5
Bare land	−0.02–0.05	<3.5
Grassland	0.05–0.1	<3.5
Tree	>0.1	>3.5

Knowledge engineer The fundamental building blocks of an expert system include hypotheses (problems), rules, and conditions. The rules and conditions operate on data (information). It is possible to address more than one hypothesis in an expert system.

The best way to conceptualize an expert system is to use decision tree structure where rules and conditions are evaluated in order to test hypotheses (Jensen 2005).

Hypotheses: The class to be tested (extracted) from the spatial data; in our case, this class is building.

Rules: A human expert should develop the knowledge base (hypotheses, rules, and conditions) to identify building from other classes. The rules and conditions were based on remote sensing multispectral reflectance and derivatives (e.g., NDVI), elevation data.

Conditions: The expert identifies very specific conditions that are associated with the remote sensing reflectance data, elevation data (Jensen 2005).

Building detection module based on multispectral classified image resulted from the third approach (building mask), NDVI, nDSM has been implemented.

Implementation of expert system The knowledge-based system was implemented to run the knowledge building detection on ERDAS Imagine software. The implemented expert system superimposes maximum likelihood classification resulted from the third approach into knowledge-based system with spectral derivative (NDVI) and height data nDSM. The knowledge-based consisted of logic or rule that determined building. Finally, the result for expert system is building class.

Since knowledge-based system inference is a way to show the relationship among data with union or mixed forms from decision tree. For building detection, building was found when nDSM is greater than 3.5 m (the height thresholds Δhmin=3.5 m that applied to nDSM) and NDVI is more than 0.038; these values have been chosen from the histogram of the NDVI, then land cover type was recognized as building.

Figure 12 shows buildings resulted from the expert system, and the four building blocks used for quality

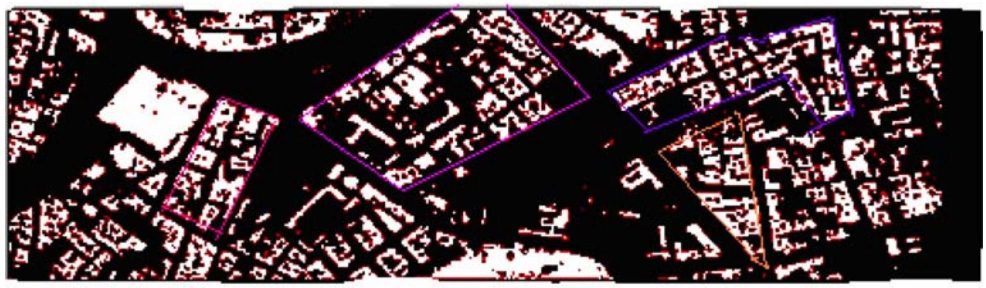

Fig. 12 Buildings resulted from the expert system. The four building blocks used for quality assessment of extraction results has been highlighted.

assessment of extraction results has been highlighted. Figure 13 shows the workflow of the expert system approach.

The overall accuracy resulted from the proposed approach was 96%, and kappa coefficient is 0.95. This indicates that the proposed approach gives better results than the three approaches.

Accuracy assessment

Classification accuracy can be determined by creating an error matrix. The error matrix consists of an $n \times n$ array, where n is equal to the number of categories or classes on the map. One axis presents the categories (classes) as derived from the remotely sensed classification, and the other axis shows the

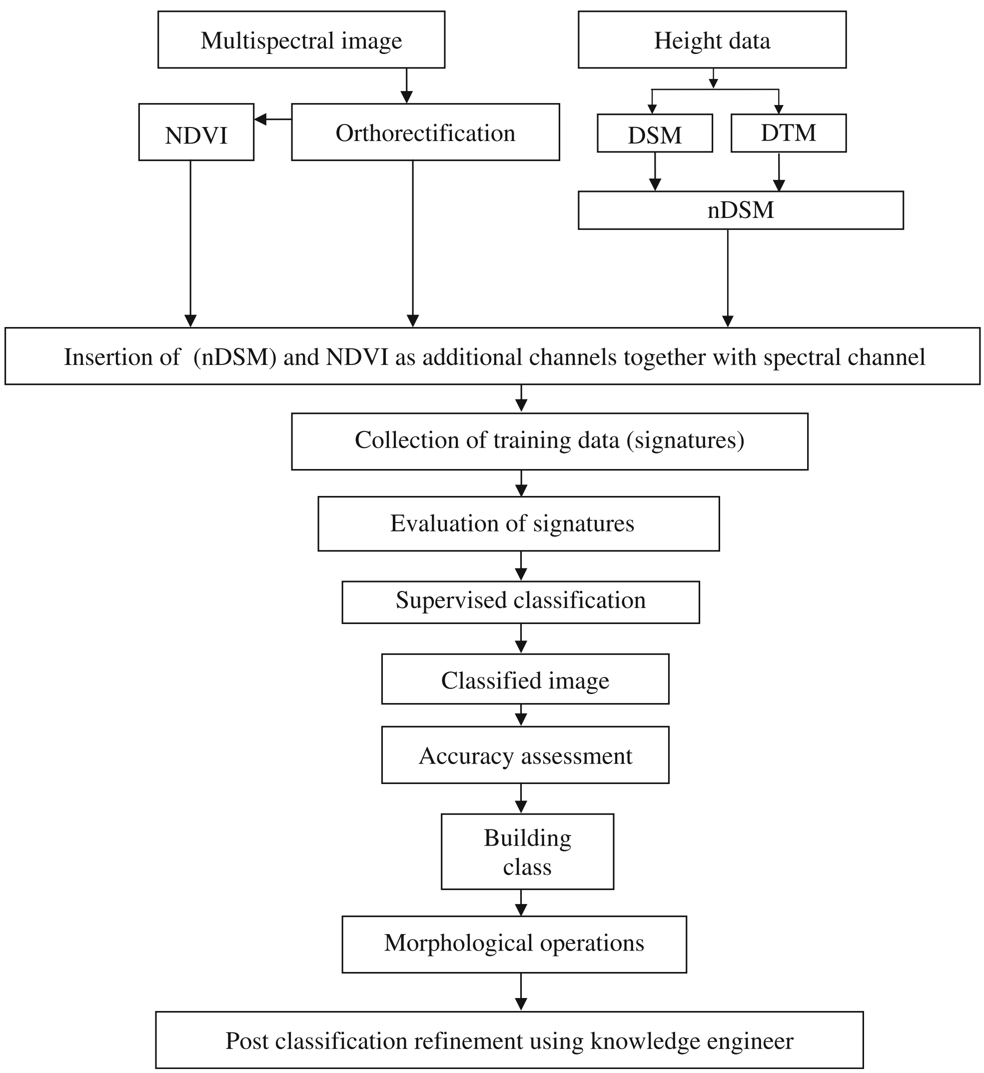

Fig. 13 Building detection from the expert system approach.

classes identified from the reference data (Macleod and Congatton 1998). In some researches, the left hand side of the matrix is labeled with the classes on the reference (correct, verified, identified, and known) or true classification; the upper edge is labeled with the same classes, but it refers to the map to be evaluated (Mohamed 1998). Some other researches take the opposite as in Janssen and Derwel (1994). Both of them are right. The diagonal from the upper left to the lower right gives correctly classified points, so the sum of these values gives the total of correctly classified points (Mohamed 1998). The overall accuracy measures the accuracy of the entire image without indicating the accuracy of the individual classification categories. It is the total number of correctly classified samples divided by the total number of reference samples.

$$O = \frac{\sum_{i=1}^{r} x_{ii}}{N}$$

Seventy randomly selected points were used to evaluate the accuracy of classification. An accurate estimation of their classes were carried out from the map and compared with the corresponding classes resulted from the classified image. Moreover, comparison on accuracy obtained from expert system and maximum likelihood classification had been done. It can be concluded that the result of the expert system provided higher overall accuracy than the maximum likelihood classification. Figure 14 illustrates the overall accuracy of the three approaches, overall accuracy of three approaches after applying morphological filters and overall accuracy of the expert system. It is clear from the figure that the overall accuracy increased from the first approach to the second to the third, and using the morphological filters increase the overall accuracy. Also, it is clear that the overall accuracy resulted from the implemented system is higher.

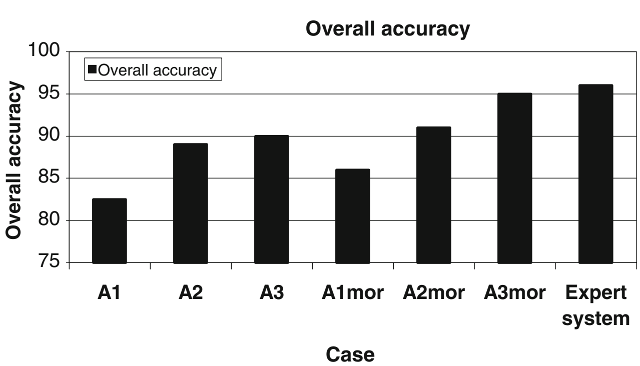

Fig. 14 Overall accuracy of the three approaches and overall accuracy of three approaches after applying morphological filters and overall accuracy of the expert system.

Kappa coefficient of agreement

The kappa coefficient of agreement is a discrete multivariate analysis technique used to evaluate the accuracy of classification maps created with remotely sensed imagery. The kappa coefficient is calculated from the error matrix and measures how the classification performs compared with the reference data. Kappa is used to determine if a classification produced from remotely sensed imagery is better than random. The kappa coefficient of agreement is the difference between the actual agreement (major diagonal total) and the chance agreement (row or column totals) of the matrix. The kappa coefficient was recommended because it considers all elements of the confusion matrix.

The kappa can be defined as:

$$K = \frac{\text{observed accuracy} - \text{chance agreement}}{1 - \text{chance agreement}}$$

and it is computed as:

$$K = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} (X_{it} \cdot X_{ti})}{N^2 - \sum_{i=1}^{r} (X_{it} \cdot X_{ti})}$$

where

- r Number of rows in error matrix
- X_{ii} Number of observations in row (i) and column (i) on the major diagonal
- X_{it} Total of observations in row i shown as marginal total of the matrix
- X_{it} Total of observations in column i shown as marginal total at the bottom of the matrix
- N Total number of observations, included in the matrix

Figure 15 shows Kappa coefficient of the three approaches, kappa coefficient of three approaches after applying morphological filters, and kappa coefficient of the expert system. It is clear from the figure that the Kappa coefficient increased from the first approach to the second to the third, and using the morphological filters increase the kappa coefficient. Also, it is clear that the Kappa coefficient resulted from the implemented system is higher.

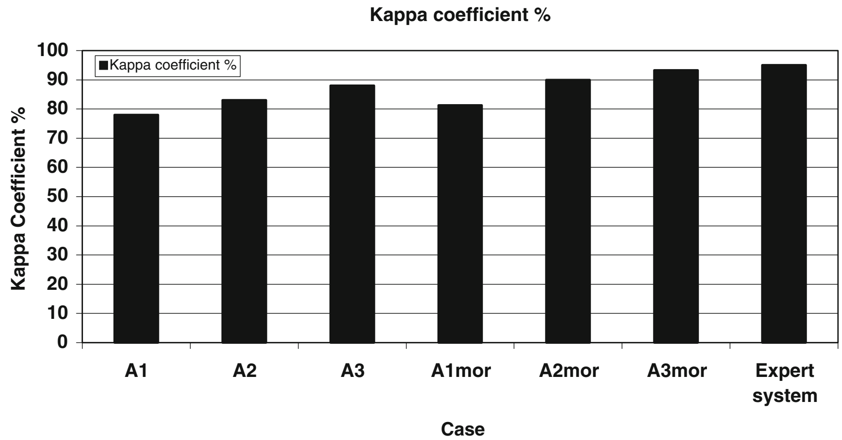

Fig. 15 Kappa coefficient of the three approaches and kappa coefficient of three approaches after applying morphological filters and kappa coefficient of the expert system.

Building extraction (vectorization)

Reference data was captured by digitizing buildings from false color composite of multispectral image with nDSM as an additional layer. Figure 16 shows buildings resulted from manual vectorization.

Quality assessment of extraction results

Comparison of building extraction results with manual on-screen digitizing vector (traditional method of extraction)

Manual on-screen digitizing from orthorectified IKONOS image has been performed. Vector results of buildings were compared with the extracted buildings from building mask resulted from expert system. Four buildings blocks have been used in this comparison. A set of indexes for comprehensively evaluating the results of automated building extraction has been used as enumerated below.

For each building block, the “branching factor”, “miss factor”, “building detection percentage”, and “quality percentage” were calculated as follows:

$$\text{Branching Factor(BF)} = \text{FP}/\text{TP}$$

$$\text{Miss Factor} = \text{FN}/\text{TP}$$

$$\text{Completeness} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{Correctness} = \text{TP}/(\text{TP} + \text{FP})$$

$$\text{Building Detection Percentage(DP)} : 100 \times \text{TP}/(\text{TP} + \text{FN})$$

$\text{Quality Percentage} : 100 \times \text{TP}/(\text{TP} + \text{FP} + \text{FN})$ (San and Turker 2007; Rottensteiner et al. 2007) where

TP Is true positive in which both the automated and manual methods classify the area as building

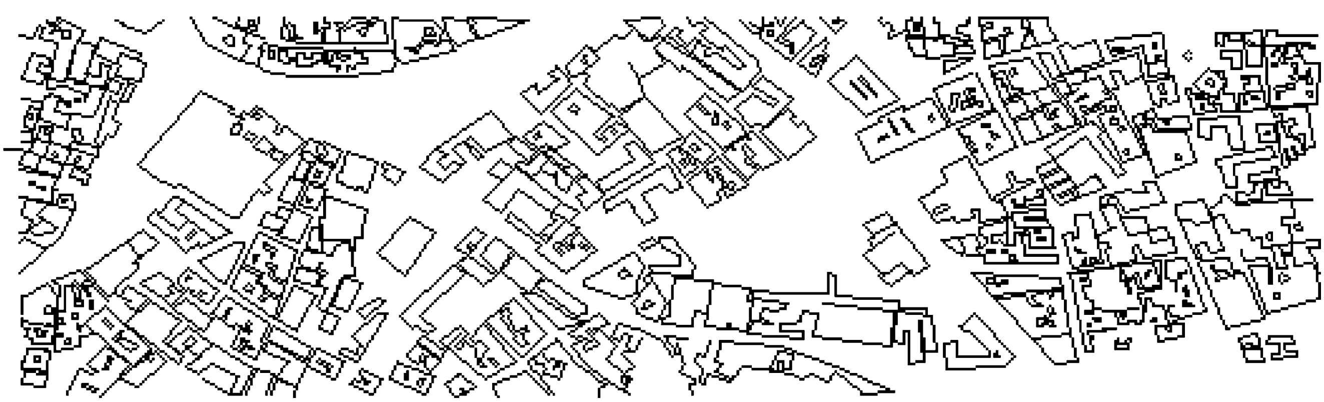

Fig. 16 Buildings resulted from manual vectorization.

TN Is true negative in which both the automated and manual methods classify the area as non-building
FP Is false positive in which only the automated method classifies the area as building
FN Is false negative in which only the manual method classifies the area as building

The "branching factor" indicates the rate of incorrectly labeled building areas, while the "miss factor" describes the rate of missed building areas. The "building detection percentage" gives the percentage of building areas correctly extracted by the automatic process, and the "quality percentage" is the overall measure of performance which accounts for all misclassifications and describes how likely a building area produced by the automatic extraction is true (San and Turker 2007; Lari and Ebadi 2007). Four building blocks of different sizes and shapes have been chosen for testing the quality of the implemented approach (see Table 3). Table 3 indicates calculation of branching factor, miss factor, completeness, correctness, building detection percentage, and quality percentage for each one of the four building blocks.

For the four building blocks, the average building detection percentage and the average quality percentage were computed to be 81.93 and 51.39, respectively.

Although the third approach appears to be quite successful especially in solving the problem of building detection for those urban blocks that contain closely located buildings as well as in separation of buildings from trees, so far, it still needs improvement.

Morphological opening with kernel size of 3×3 followed by morphological closing with kernel size of 3×3 have been applied to each of the resulted three building masks using ENVI 4.2 software in order to remove artifacts resulted in an improvement in the building detection results and the overall accuracy.

The third approach results have been improved by developing a building detection module based on integration of classified image, elevation data (LIDAR data), and spectral information derivative (NDVI).

A rule-based expert system consists of essentially hypothesis (output; buildings), and variables of a knowledge base were developed in the knowledge engineer of ERDAS Imagine for post-classification refinement of initially classified output building mask. Classification rules were enriched with ancillary data such as the nDSM and the NDVI. Each rule is a representation of each node in the tree that describes a building class or probability of presence of buildings pixel.

It has been found that the use of an expert system, which considers expert knowledge, would further help in the discrimination of the classes and improve classification accuracy of buildings. It can be concluded that the result of the expert system provided higher overall accuracy than the maximum likelihood classification; the overall accuracy of expert classification was 96%, and kappa coefficient was 0.95.

Table 3 Quality assessment of the results of building extraction from expert system.

Quality assessment	Building block1	Building block2	Building block3	Building block4
FP	2	11	0	4
FN	0	1	8	0
TP	1	17	4	7
MF	0	0	2	0
BF	2	0.647	0	0.571
DP	100	94.4	33.33	100
QP	33.33	58.62	50	63.636
COMP	1	0.944	0.333	1
CORR	0.333	0.6071	1	0.636

Conclusions

Three approaches for building detection from IKONOS image based on maximum likelihood classification have been compared. The results show that the third approach is the best for building detection followed by the second approach then the first approach. Some buildings have some parts missing (not completely detected); this can be referred to the detection rate decreases with small building size.

After that, rectified multispectral QuickBird image obtained from Google earth has been used for revision of the resulted buildings specially buildings that are not discernible enough due to shadows of IKONOS image. This step was used instead of the fieldwork.

The resulted vector map of buildings from on-screen digitizing of false color composite of multispectral image with nDSM as an additional layer can be considered as a starting point for further algorithm developments.

For the four building blocks that were used for assessment of the quality of extraction results, the average building detection percentage and the average quality percentage were computed to be 81.93 and 51.39, respectively.

It was found the 1:5,000 map obtained from the ESA doesn't show each building separately but as a building block. The implemented method is quite successful for obtaining the same result, and it can give details inside each building block.

It is recommended to

- Use the shape cue for building extraction in order to improve the results.
- Consider the manual vectorization of buildings from false color composite of multispectral image with nDSM as an additional layer as a starting point for further algorithm developments.

Also, it is recommended to do additional researches:

- To use object-based classification.
- To fuse LIDAR data and multispectral images in order to improve building detection.

Acknowledgment The authors thank NARSS for giving us the data. The editing and comments of the reviewers is gratefully appreciated.

References

Baltsavias EP (2002) Object extraction and revision by image analysis using existing geospatial data and knowledge: state-of-the-art and steps towards operational systems. International Archives of Photogrammetry and Remote Sensing, Commission II, IC WG II/IV, http://e-collection.ethbib.ethz.ch/ecol-pool/bericht/bericht_345.pdf

Forghani A (1999) An expert system approach for detection of roads from remote sensing data. For presentation at the Joint Workshop of ISPRS Working Groups 1/1, 1/3, and IV/4: Sensors and Mapping from Space, Honover, Germany, 27–30 September 1999

Guo T, Yasuoka Y (2002a) Combining high resolution satellite imagery and airborne laser scanning data for generating Bareland Dem in urban areas. Proceedings of International Workshop on Visualization and animation of landscape, acrors.ait.ac.th, http://www.commission5.isprs.org/kunming02/download/GuoTao.pdf

Guo T, Yasuoka Y (2002b) Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas. www.gisdevelopment.net/aars/acrs/2002/vhr/018.pdf

Haala N (1999) Combining multiple data sources for urban data acquisition. photogrammetric week 1999. www.ifp.uni-stuttgart.de/publications/phowo99/haala99.pdf

Haala N, Walter V (1999) Aerial automatic classification of urban environments for database revision using LIDAR and color imagery. International Archives of Photogrammetry and Remote Sensing 32(Part 7-4-3):W6, Valladolid, Spain, 3–4 June, http://www.data-fusion.org/ps/sig/meeting/Spain99ps/haala.pdf

Jensen JR (2005) Introductory digital image processing a remote sensing perspective, 3rd edn. Pearson Prentice Hall, Upper Saddle River

Janssen LF, Derwel FJMV (1994) Accuracy assessment of satellite derived land cover data: a review. Photogrammetric Engineering and Remote Sensing 60(4):419–426

Lari Z, Ebadi H (2007) Automatic extraction of building features from high resolution satellite images using artificial neural networks. ISPRS 2007, www.isprs2007ist.itu.edu.tr/25.pdf

Lu YH, Trinder J, Kubi K (2003) Automatic building extraction for 3D terrain reconstruction using interpretation techniques. http://www.ipi.uni-hannover.de/html/publikationen/2003/workshop/yihuilu.pdf

Macleod RD, Congatton RG (1998) A quantitative comparison of change detection algorithms for monitoring Eelgrass from remotely sensed data. Photogrammetric Engineering and Remote Sensing 64(3):207–216

Marangoz AM, Alkis Z, Karakis S (2007) Evaluation of information content and feature extraction capability of very high resolution pan-sharpened QuickBird image. Commission VII, WG2 & WG7 ISPRS www.isprs2007ist.itu.edu.tr/34.pdf

Mohamed A (1998) Assessment of classified remote sensing imagery using various techniques. M.SC. Thesis, Cairo University, Cairo

Muchoney D, Borak J, Chi H, Friedl M, Gopal S, Hodges J, Morrow N, Strahler AH (2000) Application of MODIS global supervised classification model to vegetation and land cover mapping in Central America. International Journal of Remote Sensing 21:115–1138

Nangendo G, Skidmore AK, Oosten HV (2007) Mapping East African tropical forests and woodlands—a comparison of classifiers. ISPRS Journal of Photogrammetric Engineering and Remote Snesing 61(6):393–404

Pal M, Mather PM (2003) An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of the Environment 86:554–565

Rottensteiner F, Briese C (2002) A new method for building extraction in urban areas from high-resolution LIDAR data. ISPRS 2002 http://www.isprs.org/commission3/proceedings02/papers/paper082.pdf

San KD, Turker M (2007) Automatic building extraction from high resolution stereo satellite images. Commission VII ISPRS 2007 http://www.isprs2007ist.itu.edu.tr/39.pdf

Tso B, Mather PM (2001) Classification methods for remotely sensed data. Taylor & Francis, New York